THE KINETICS OF THE REACTION OF SUBTILISIN BPN' WITH CHLOROMETHYL KETONES IN RELATION TO ITS SUBSITE SPECIFICITY

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

James T. Tippett

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Chemistry

Georgia Institute of Technology

March, 1973

THE KINETICS OF THE REACTION
OF SUBTILISIN BPN' WITH CHLOROMETHYL KETONES
IN RELATION TO ITS SUBSITE SPECIFICITY

Approved:

James C. Powers, Chairman

Drury D. Caine, III

Charles L. Liotta

Date approved by Chairman: 2/26/73

ACKNOWLEDGMENTS

The author is deeply grateful to Dr. James C. Powers for his invaluable guidance in accomplishing the work presented in this thesis and for providing financial support through grants from the National Institutes of Health.

The reading of this thesis by Dr. Drury S. Caine, III and Dr. Charles L. Liotta is greatly appreciated.

Also, I would like to thank my parents, Dr. and Mrs. John Tippett, Jr., for their constant encouragement and inspiration.

Finally, special appreciation is due Miss Sandra Hayles for her assistance in organizing and typing this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY OF ABBREVIATIONS

Abbreviation	Correct Terminology
BOC	tert-butyloxycarbonyl
C	centigrade
d	doublet
g	gram(s)
H	proton(s)
m	multiplet
mm	millimeter
mM	millimolar
M	molarity
m/e	mass to charge ratio
mg	milligram(s)
ml	milliliter(s)
mp	melting point
nmr	nuclear magnetic resonance
p	parent ion
psi	pounds per square inch
s	singlet
t	triplet
Tos	tosyl
Z	benzyloxycarbonyl

SUMMARY

The crystallographic model of the enzyme subtilisin inhibited with peptide chloromethyl ketones has verified the existence of a substrate binding site composed of well-defined subsites. Based on these findings and the interactions observed, rationalization of previous kinetic data on true substrates is possible. However, these rationalizations are valid only if the assumption is made that the binding of inhibitors in the crystallographic model is a good model for substrate binding. Therefore, the purpose of this work was to determine the rates of reaction of a series of peptide chloromethyl ketones with subtilisin and to correlate these rates with the number of interactions observed in the crystallographic model and the known specificities of the subsites observed from the rates of substrate hydrolysis. The positive correlations obtained and presented in this thesis provide convincing evidence for the postulation that the productive binding for these inhibitors which leads to alkylation of His-64 is closely related to the productive binding modes of peptide substrates which leads to the acylation of Ser-221.

CHAPTER I

INTRODUCTION

Subtilisin is a bacterial enzyme discovered by Linderstrøn-Lang and Ottesen (1947) during an investigation of the conversion of ovalbumin to plakalbumin. This enzyme is one member of a widely distributed family of enzymes called proteases or proteolytic enzymes which are required in biological systems to catalyze the hydrolytic cleavage of peptides. Subsequent research showed that there were a series of subtilisin proteases exhibiting similar chemical properties but varying amino acid composition. The enzyme isolated from Bacillus amyloliquefacians, designated Subtilisin BPN', has been most extensively studied and in this thesis will be referred to simply as subtilisin. Other members include Subtilisin Carlsberg, Novo, and Amylosacharriticus, but it has recently been shown that the Novo enzyme is identical to BPN' in all respects (Kraut, 1971).

Proteases are divided into four classes (Hartley, 1960): metallopeptidases, sulfydryl proteases, acid proteases, and serine proteases. The latter group of which subtilisin is a member, is classed together due to a common functional feature - the presence of a unique, highly reactive serine residue at the active site. The reagent

diisopropylfluorophosphate (DFP) selectively phosphorylates this residue despite the presence of numerous other serine residues in the enzyme, and the resulting DIP derivative shows no enzymatic activity. Matsubara et al. (1958) completely inhibited subtilisin with DFP, crystallized the DIP derivative, and showed that one mole of the inhibitor was bound to a serine residue per mole of enzyme. Subsequently, Noller and Bernhard (1965) used a specific chromophoric acylating agent, furylacryloylimidazole, to label and identify the reactive serine of subtilisin. A furylacryl peptide was isolated after enzymatic digestion, and a sequence determination showed the labeled residue to be Ser-221.

Another characteristic enzymatic property of serine proteases is the presence of a unique histidine residue. Oosterbahn and Cohen (1964) provided the first evidence for histidine involvement in subtilisin by photooxidation studies. Correlation of the loss of a histidine residue and enzymatic activity implicated histidine involvement in the active site. In addition, kinetic data by Polgar and Bender (1967) showed that decreases in subtilisin activity was concomitant with protonation of a group with a pK of 7.2; this group was presumed to be the imidiazole of histidine. Although the above evidence suggested the involvement of a histidine residue, direct chemical evidence was lacking. Specific reagents used for labeling the active site histidine

residue of other serine proteases were unreactive with subtilisin (Smith et al. 1966). However, Shaw and Ruscica (1968) were able to synthesize a related active site specific reagent which was highly reactive with subtilisin. The new reagent, benzyloxycarbonyl-L-phenylalanine bromomethyl ketone (Z-PheCH_2Br), was shown to stoichiometrically alkylate a single histidine residue, and loss of enzymatic activity was demonstrated by the finding that the alkylated derivative no longer reacted with DFP. Markland et al. (1968) used tritium labeled Z-PheCH_2Br to identify the reactive histidine. After suitable degradation of the labeled enzyme by proteolytic hydrolysis and cyanogen bromide cleavage, peptides were isolated that were shown to contain histidine residue 64 substituted in the three position of the imidazole ring.

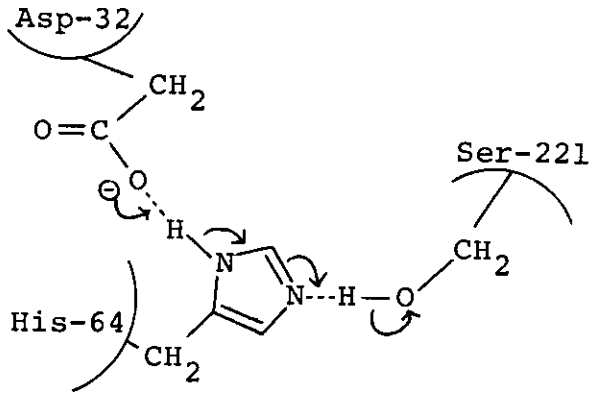
From the kinetic and chemical modification studies described above, both the histidine residue at position 64 and the serine residue at position 221 of subtilisin have been identified as participating in catalysis even though the two residues are remote from one another in the linear primary amino acid sequence. It was therefore predicted that these residues must come in close proximity in the tertiary structure of the enzyme forming part of the catalytic site. High resolution x-ray diffraction studies have indeed shown that a hydrogen bond exists between these two residues forming part of a hydrogen bond network at the

catalytic site (Alden et al., 1970). Contained in this network is a hydrogen bond from the OH group of Ser-221 to $N^{\varepsilon}2$ of His-64; a hydrogen bond from $N^{\varepsilon}1$ of His-64 to the side chain $O^{\delta}2$ of Asp-32; and another hydrogen bond which extends to $O^{\delta}2$ of Asp-32 from the side chain OH group of Ser-33. All atoms involved in this network (1) lie approximately within the plane of the imidazole ring. This network makes possible a charge relay system whereby the negative charge on the aspartate residue 32, buried in a hydrophobic environment in the enzyme, may be transferred to the surface of the molecule at Ser-221. This Ser-221 exists as a highly nucleophilic alkoxide group in one of the canonical forms (2) of the system. This system has also been observed in other serine proteases (Birktoft et al., 1970).

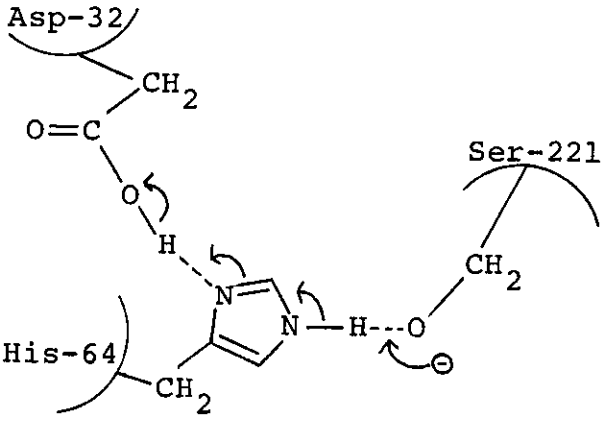
The finding that the ultimate site of acylation in serine proteases is the hydroxyl oxygen of a unique serine residue, led to the hypothesis that enzymatic hydrolysis might occur via an intermediate acyl enzyme. However, since these results were previously established by chemical analysis of the degraded enzyme, it is possible that acylation initially occurred elsewhere in the enzyme and was transferred to serine during the degradation. Therefore, Bernhard et al. (1965), sought to confirm the hypothesis by observing an acyl intermediate of subtilisin in solution. Comparison of the ultraviolet (uv) spectra of chromophoric acyl enzymes with

presumed O-acylserine peptide analogues, showed that the cinnamoyl- and furylacrylolyl- proteases, isolated at low pH, have UV absorption maxima characteristic of an oxygen ester. Subsequently, Polgar and Bender (1966) converted the active serine of subtilisin into a cysteine residue. Despite decreased protease activity, the enzyme was cinnamocylated with cinnamoylimidazole and the intermediate showed ultraviolet absorption maxima characteristic of a thiol ester. These results supported the hypothesis that peptide bond cleavage might occur via a two step process involving acylation and deacylation. A possible mechanism incorporating these results has been proposed and is illustrated below (Blow et al., 1969). Following binding of the substrate to the enzyme forming the enzyme-substrate couplex (3), the carbonyl carbon of the substrate is attacked by the serine oxygen forming the acyl enzyme (4). Deacylation of this intermediate is presumed to be the exact reverse of the acylation step, except with water in the position formerly occupied by the leaving group (5). Thus, the enzyme is restored to its active form for further catalysis (6). The crystallographic model of the active site of subtilisin is shown in Figure 1.

As previously mentioned, enzymes exhibit substrate specificity primarily in the types of chemical bonds that they cleave and are often classed accordingly. A variety of large proteins have been hydrolyzed by subtilisin

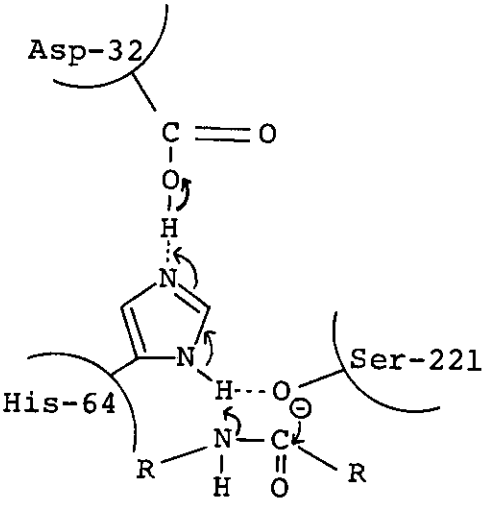

Asp-32
CH2
O=C
O
H
N
N---H—O
Ser-221
CH2
His-64
CH2

(1)

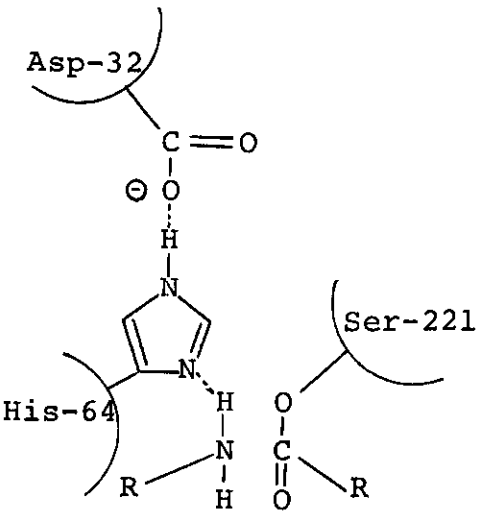

Asp-32
CH2
O=C
O
H
N
N—H---O
Ser-221
CH2
His-64
CH2

(2)

Acylation Step

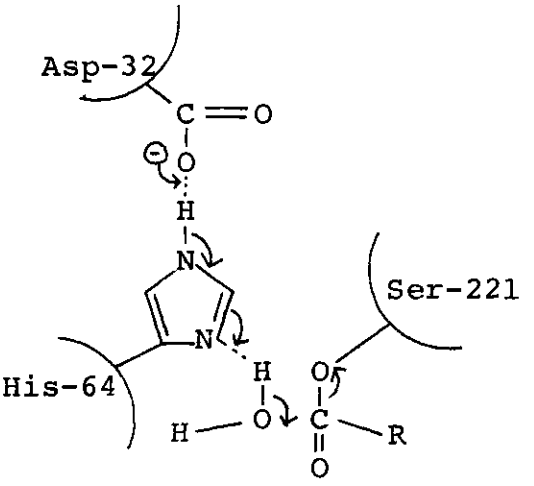


(3)

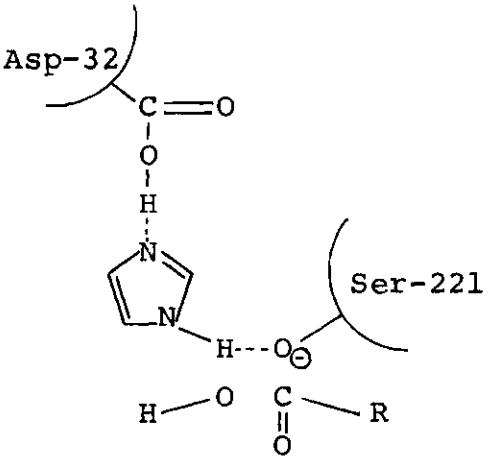


(4)

Deacylation Step



(5)



(6)

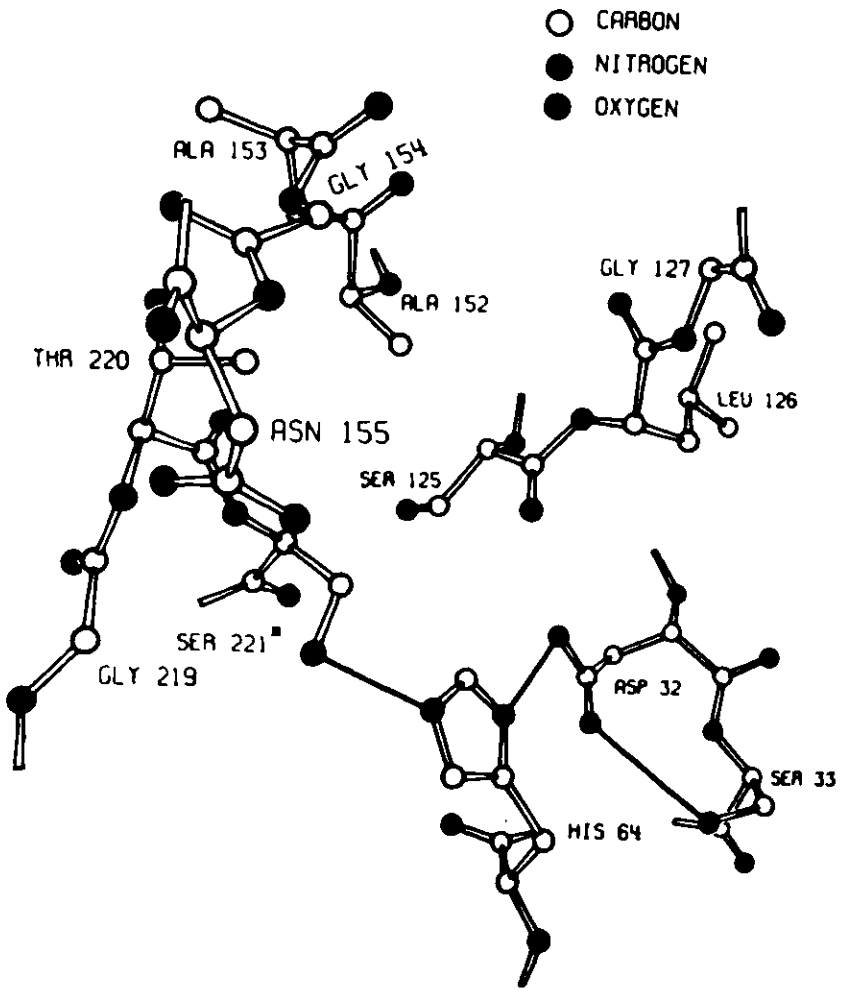


Figure 1. Crystallographic Model of the Active Site of Subtilisin.

including casein, hemoglobin, gelatin, oxidized lysozyme, and insulin (Okunuki et al., 1956; Tuppy, 1953). In each case, the enzyme cleaved not only at internal peptide bonds but also at terminal bonds yielding free amino acids. Due to this activity on peptide bonds subtilisin was classed as a protease and more specifically as an endopeptidase, splitting peptide linkages at any point in a peptide chain (as opposed to an exopeptidase which can split only terminal peptide bonds). In addition to being a protease, subtilisin was also shown to have esterase activity, hydrolyzing small synthetic esters (Matsubara et al., 1958). From many early studies, however, it has been determined that enzymes exhibit not only primary specificity for a particular chemical bond, but also a secondary specificity for another functional group (or groups) in the substrate essential for its binding to the enzyme such that the sensitive bond is properly oriented with respect to the catalytic groups (Lehninger, 1970). Consequently, subtilisin would be expected to cleave peptide bonds in proteins, but not indiscriminately. Morihara and Tsuzuki (1969) reinvestigated the subtilisin catalysed cleavage of insulin under restricted conditions and found secondary specificity in its discriminate cleavage at the carbonyl group of amino acid residues with aromatic or apolar side chains. These observations were consistent with the subsequent finding of these authors that hydrolysis of small amide and ester substrates such as CBZ-Gly-X-NH_2

and acetyl-X-ethyl ester occurred at the carbonyl-terminal side of residue X, where X represented such residues as L-tyrosine, L-phenylalanine, and L-leucine. To further study these secondary specificities, Morihara et al. (1969) synthesized a series of oligopeptides and examined the effect of neighboring residues surronding the sensitive linkage. An acceleration in the rate of hydrolysis was observed when the peptide chain was elongated on the N-terminal side of the point of catalytic cleavage: CBZ-Leu-$\underset{\uparrow}{NH_2}$ (0.23) CBZ-Gly-Leu-$\underset{\uparrow}{NH_2}$ (1) CBZ-Gly-Gly-Leu-$\underset{\uparrow}{NH_2}$ (36.1) (the arrow indicates the point of cleavage and the numbers in parentheses are relative rates of hydrolysis). It was also found that hydrolysis is inhibited when the terminal α-amino and α-carbonyl groups of these substrates were unblocked, if the former group within four residues on the N-terminal side and the latter group within one residue on the C-terminal side of the cleavage point. Thus, it is postulated that the active site of subtilisin, and other serine proteases is composed of two specific zones. One is the catalytic zone (or catalytic site) composed of the residues in the enzyme immediately responsible for bond breaking, and the second is a binding site where the substrate is bound to the enzyme. The binding site appears to be composed of a sequence of subsites. A scheme for depicting the positioning of a polypeptide chain in the binding site of a protease has been proposed (Schecter and Berger, 1967) and is shown

in Figure 2.

Morihara et al. (1970) found that in subtilisin each subsite exhibits individual stereo- and side chain specificities. Stereospecificity was exemplified by the fact that the site of hydrolysis of polypeptide amides was greatly reduced when D-amino acid residues were substituted for the corresponding diasterioisomers. In fact, a distinct preference was shown for L-amino acids at all subsites. Hydrolysis rates also showed dependence on the particular side chain at a given subsite; for example, alanine at S_2 greatly promoted activity when compared with glycine, histidine, or tyrosine.

To account for the aforementioned specificities and to determine the specific interactions occurring at the individual subsite, crystallographic structures of substrates actually bound to the enzyme must be observed. However, enzymatic reactions occur on a time scale in the neighborhood of 10^{-3} seconds or faster and collection of adequate x-ray data requires days. Robertus et al. (1971) circumvented this problem by the use of irreversible covalent inhibitors believed to be analogues of good polypeptide substrates. Four polypeptide chloromethyl ketone inhibitors with an L-phenylalanine residue at their C-termini were bound to subtilisin and investigated. All were found to alkylate the catalytic site His-64 at $N^{\varepsilon}2$ with their phenylalanine side chain fitting into a hydrophobic crevice. The

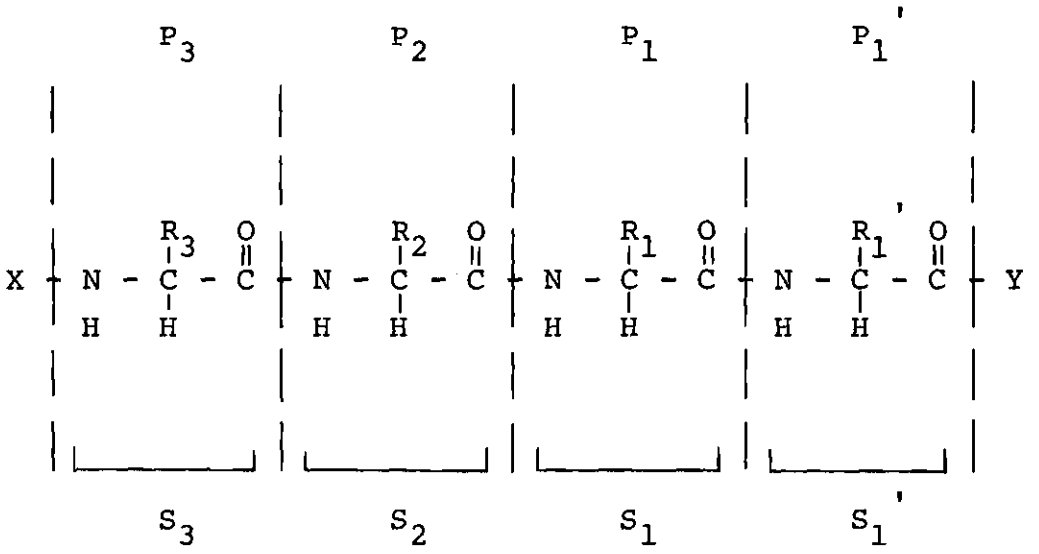


Figure 2. Scheme for Depicting the Positioning of a Polypeptide Chain in the Binding Site of a Protease. (Schecter and Berger, 1967)

Amino acid residues P_3, P_2, P_1 and P_1' are located at subsites S_3, S_2, S_1 and S_1', respectively. Side chains are designated by R. Cleavage occurs between residues P_1 and P_1'. Additional subsites exist at either end, X or Y.

polypeptide chain of each inhibitor was extended and bound in the manner of an antiparallel β-sheet to a segment of extended backbone chain in the enzyme consisting of Ser-125 - Leu-126 - Gly-127. A detailed description of the interaction at each subsite was obtained. These interactions and the mode of binding are illustrated in Figure 3 and will be described briefly.

At subsite S_1, the phenylalanine side chain of P_1 lies in a crevice formed by three pieces of the enzyme molecule. Two sides of the crevice are flat and well-defined, but the third is less regular. A flat face of the crevice consists of the extended backbone of segment Ser-125 - Leu-126 - Gly-127. A second interaction at S_1 is a hydrogen bond between the NH of P_1 and the carbonyl oxygen of Ser-125.

Interactions at S_2 are less well-defined. The NH and CO of P_2 point toward the solvent, but interaction of the side chain with residues in the enzyme appears likely.

Interactions between P_3 and S_3 are again quite well-defined. Hydrogen bonds are possible between the CO of P_3 and the backbone NH of Gly-127, and between the NH of P_3 and the backbone CO of Gly-127. The side chain of P_3 appears not to interact with subsite S_3.

The nature of the possible interactions at S_4 was not observable since only tripeptide inhibitors were studied.

From the crystallographic model of the inhibited enzyme, the existence of a binding site composed of well-

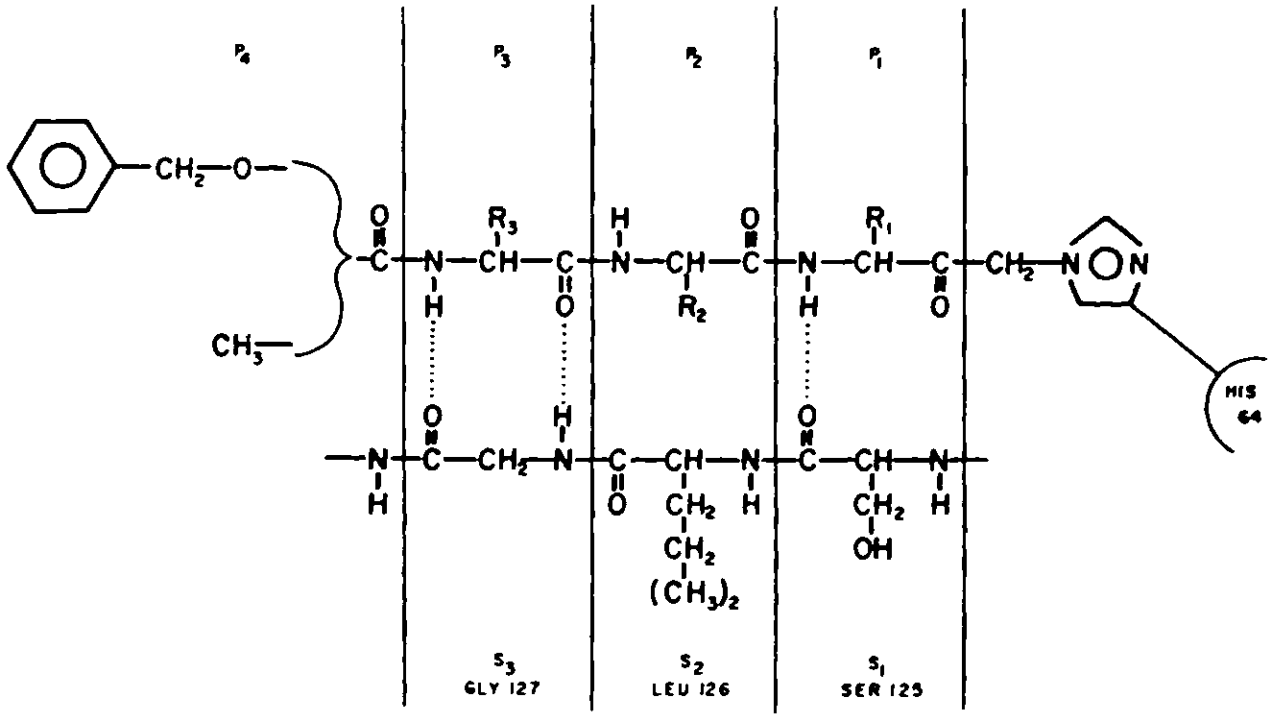


Figure 3. Mode of Binding Peptide Chloromethyl Ketone Inhibitors to Subtilisin.

defined subsites has been verified. Based on the findings and the interactions observed, rationalization of previous kinetic data on true substrates is possible. However, these rationalizations are valid only if the assumption is made that the binding of inhibitors in the crystallographic model is a good model for substrate binding. Therefore, the purpose of the work presented in this thesis was to determine the rates of reaction of a series of peptide chloromethyl ketones with subtilisin and to correlate these rates with the number of interactions observed in the crystallographic model and the known specificities of the subsites observed from the rates of substrate hydrolysis. Positive correlations would provide convincing evidence for the postulation that the productive binding mode for these inhibitors which leads to alkylation of His-64 is closely related to the productive binding modes of peptide substrates which leads to the acylation of Ser-221. A second objective was to synthesize a tetrapeptide leucine chloromethyl ketone for future crystallographic studies of the binding interactions at the fourth and first subsites respectively.

CHAPTER II

EXPERIMENTAL

Methods and Materials

The following reagents were obtained from commercial sources: optically active amino acids from Ajinomoto Co., Inc., Japan (all amino acids used were of the L form unless otherwise stated); glycylglycine from Pierce Chemical Co.; benzylchloroformate and tert-butyl azidoformate from PCR, Inc.; i-butyl chloroformate and i-valeryl chloride from Eastman Chemical Co.; N-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ) from Aldrich Chemical Co.; and Subtilisin BPN' from Nagase and Co., Japan (used without further purification).

The enzyme substrate N-acetyl-L-tyrosine ethyl ester monohydrate was synthesized in the laboratory by standard methods and recrystallized from ethanol-water to constant melting point (79.5°-80.5°). Anhydrous magnesium sulfate was employed in drying organic solutions and extracts in organic solvents.

Tetrahydrofuran (THF) was dried by refluxing 24 hours over sodium metal and then distilling from benzophenone ketyl. Dimethylformamide (DMF) was dried for 48 hours over 4A molecular sieves. Anhydrous N-methyl morpholine was

obtained by distillation from potassium hydroxide pellets. All other common chemicals and solvents were analytical grade.

Mass spectra were obtained using a Varian Model A-66 medium resolution mass spectrometer. Nuclear magnetic resonance (nmr) spectra were acquired using a Varian Model A-60 spectrometer; solvents used were deuterochloroform ($CDCl_3$) and deuterodimethylsulfoxide (DMSO-d_6) containing one percent tetramethylsilane (TMS) as an internal standard, or deuterium oxide containing sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as an internal standard.

Solvent evaporation under reduced pressure was accomplished with a Buchler flash evaporator. Melting points were determined using a Buchi capillary melting point apparatus and are uncorrected. Elemental microanalyses were performed by Atlantic Microlab, Inc., Atlanta, Georgia.

All compounds isolated were shown to be pure by thin-layer chromatography. The absorbent used was silica gel G, pre-coated (Merck) on glass plates. After development in the appropriate solvent, the spots were detected in a closed chamber saturated with iodine vapor.

Peptide chloromethyl ketones in addition to those described in the experimental section were used for kinetic studies. These were previously prepared as described by Powers and Wilcox (1970) and Segal et al. (1971).

Experimental Procedures

Preparation of Inhibitors

Acetylglycylglycyl-L-alanine Benzyl Ester (Ac-Gly-Gly-Ala-OBzl). The p-toluenesulfonic acid salt of L-alanine benzyl ester was prepared by the azeotropic distillation of L-alanine, p-toluenesulfonic acid, benzyl alcohol, and benzene (Greenstein et al., 1961) and neutralized in cold sodium carbonate. Following extraction with ether and concentration under vacuum, the liquid L-alanine benzyl ester was obtained. A solution of Ac-Gly-Gly-OH (6.0 g, 34.5 mmoles; prepared by acylation of glycylglycine with one equivalent of acetic anhydride in boiling glacial acetic acid) in DMF (150 ml) was cooled in a moisture-free reaction vessel to -15°C with a dry ice-acetone bath. While stirring, N-methylmorpholine (3.79 ml, 34.5 mmoles) was added to the solution followed by isovaleryl chloride (4.16 g, 34.5 mmoles). A precipitate immediately formed and the reaction mixture was allowed to stir ten minutes. Alanine benzyl ester (6.17 g, 34.5 mmoles) was added and the flask allowed to come to room temperature. The N-methylmorpholine hydrochloride was filtered off, and the filtrate was concentrated under vacuum giving a yellow residue. The residue was washed with two 25 ml portions of water giving immediate crystallization of the product. The product was recrystallized from hot methanol yielding 6.92 g (60%): mp 191-192°; nmr (DMSO-d_6) showed signals at δ 1.32 (3H, d, alanine methyl), 1.88 (3H,

s, acetyl), 5.13 (2H, s, C_6H_5-$\underline{CH_2}$-), 7.35 (5H, s, C_6H_5-); mass spectrum m/e (relative intensity) 263 (7, P-$CH_3CONHCH_2$), 221 (5), 149 (9), 128 (66, $NHCH_2CONHCH(CH_3)CO$), 107 (16, $C_6H_5CH_2O$), 91 (100, $C_6H_5CH_2$), 85 (28, $COCH(CH_3)NHCO$), 57 (50, $NHCH_2CO$).

Anal. Calcd for $C_{16}H_{21}N_3O_5$: C, 57.31; H, 6.26; N, 12.53. Found: C, 57.34; H, 6.40; N, 12.61.

The above compound was also obtained by activation of Ac-Gly-Gly-OH (4.25 g, 24.4 mmoles) with EEDQ (6.18 g, 25 mmoles) in DMF (200 ml). Alanine benzyl ester (4.37 g, 24.4 mmoles) was added, and the solution was stirred for 30 hours at room temperature. The solvent was removed in vacuo giving a yellow residue. Upon washing with chloroform, the yellow color was lost and white crystalline product remained. The product was recrystallized as above yielding 3.62 g (44%).

Acetylglycylglycyl-L-alanine (Ac-Gly-Gly-Ala-OH). A suspension of Ac-Gly-Gly-Ala-OBzl (1.6 g, 5.0 mmoles) and 10 percent palladium on charcoal (0.3 g) in acetic acid (150 ml) was hydrogenated on a Paar bomb under 40 psi pressure. After two hours, the catalyst was filtered off, and the filtrate evaporated in vacuo to give white crystalline product. The product was recrystallized from ethanol-ether to yield 1.10 g (90%): mp 197-199°; nmr (D_2O) showed signals at δ 1.42 (3H, d, alanine methyl), 2.08 (3H, s, acetyl), 4.00 (4H, s, glycine methylenes); mass spectrum m/e (relative intensity) 220 (13), 205 (53), 149 (53), 110 (100), 73 (80,

$CH_3CHCOOH$), 57 (53, $NHCH_2CO$), 43 (86, CH_3CO).

Acetylglycylglycyl-L-alanyl-L-phenylalanyl Chloromethyl Ketone (Ac-Gly-Gly-Ala-PheCH_2Cl). The mixed anhydride formed from Ac-Gly-Gly-Ala-OH (1.0 g, 4.0 mmoles), N-methylmorpholine (.44 ml, 4.0 mmoles), and i-butylchloroformate (.52 ml, 4.0 mmoles) in DMF (30 ml) was reacted with L-phenylalanine chloromethyl ketone hydrobromide (Segal et al., 1971) and triethylamine (.67 ml, 4.5 mmoles) for two hours. The DMF was removed under vacuum and 10 ml of chloroform was added to the residue giving immediate crystallization of the product. The product was filtered and successively washed with 1 M citric acid, 10% sodium bicarbonate, and distilled water. The resulting crystals were recrystallized from methanol-water yielding .92 g (54%): mp 222-224°; nmr (DMSO-d_6) showed signals at δ 1.14 (3H, d, alanine methyl), 1.88 (3H, s, acetyl), 3.08 (2H, d, C_6H_5-CH_2-), 3.66-3.81 (4H, m, glycine methylenes), 4.50 (2H, d, -CH_2-Cl), 7.26 (5H, s, C_6H_5-); mass spectrum m/e (relative intensity) 424 (2, P), 360 (4, P-CH_2Cl + CH_3), 317 (4, 360-CH_2CO), 200 (12, $CH_3CONHCH_2CONHCH_2CONHCHCH_3$), 180 (18, $C_6H_5CH{=}CHCOCH_2Cl$), 129 (13, $CH_3CONHCH_2CONHCH_2$), 91 (36, $C_6H_5CH_2$), 72 (74, $CH_3CONHCH_2$), 43 (100, CH_3CO).

Anal. Calcd for $C_{19}H_{25}ClN_4O_5$: C, 53.64; H, 5.88; N, 13.17; Cl, 8.47. Found: C, 53.58; H, 5.92; N, 13.09; Cl, 8.50.

N-Benzyloxycarbonyl-L-leucyl Chloromethyl Ketone

(Z-LeuCH$_2$Cl). A solution of Z-Leu-OH (3.3 g, 12.5 mmoles; prepared by the method of Schotten and Bauman from L-leucine, benzyloxycarbonyl chloride, and sodium hydroxide) was dissolved in THF (60 ml) and added to a reaction vessel over a fritted filter which was kept at a temperature of -5°C by a constant temperature circulator. While stirring, N-methylmorpholine (1.38 ml, 12.5 mmoles) was added, followed by i-butylchloroformate (1.63 ml, 12.5 mmoles). A white precipitate immediately formed; the mixture was stirred for 15 minutes and then filtered into a moisture-free flask. To the filtrate was added 25 millimoles of diazomethane in ether (Moore and Reed, 1961). After standing in the refrigerator overnight, the solution was saturated with HCl gas. The solvents were removed under vacuum giving a yellow oil. The oil was taken up in ethyl acetate and successively washed with 1 M citric acid, 10% sodium carbonate, and distilled water. After drying and removal of the solvent, 2.5 g (69%) of product was obtained as an oil: nmr ($CDCl_3$) showed signals at δ .91 (6H, d, isopropyl), 4.21 (2H, s, CH_2Cl), 5.08 (2H, s, C_6H_5-CH_2-), 7.30 (5H, s, C_6H_5); mass spectrum m/e (relative intensity) 261 (2, P-HCl), 220 (27, P-C_6H_5), 219 (9, $C_6H_5CH_2OCONHCH(CH_2)COCH_2$), 205 (2, 219-$CH_2$), 151 (4, P-$(CH_3)_2CHCH{=}CHCOCH_2Cl$), 111 (2, $(CH_3)_2CHCH{=}COCH_2$), 107 (23, $C_6H_5CH_2O$), 91 (100, $C_6H_5CH_2$), 77 (29, C_6H_5), 57 (32, $CH_2CH(CH_3)_2$), 43 (32, $CH(CH_3)_2$).

L-Leucyl Chloromethyl Ketone Hydrochloride (LeuCH$_2$Cl-

HCl). The monohydrate of BOC-Leu-OH was prepared from L-leucine, tert-butyl azidoformate, and sodium hydroxide in dioxane and water (Schnabel, 1967). The crystalline monohydrate was taken up in ethyl acetate and dried over magnesium sulfate. After filtration and removal of the solvent, a clear oil was obtained. The chloroketone of BOC-Leu-OH was prepared by a procedure analogous to that used to prepare Z-LeuCH_2Cl. However, upon saturation of the intermediate diazoketone with hydrogen chloride, the tert-butyloxycarbonyl group was removed with simultaneous formation of the chloroketone giving crystalline L-leucine chloromethyl ketone hydrochloride. Following filtration, the crystalline product was obtained in 71% yield: mp 150-151°.

Acetyl-L-alanyl-L-leucyl Chloromethyl Ketone (Ac-Ala-LeuCH_2Cl). Acylation of L-alanine was carried out in boiling acetic acid with one equivalent of acetic anhydride. N-Acetyl-L-alanine was coupled with L-leucine chloromethyl ketone hydrobromide via the mixed anhydride procedure previously described. The crude oil obtained was placed on a chromatography column of 0.05-0.02 mm Silica Gel G and the product was eluted with 4% methanol in chloroform. Crystallization from ethyl acetate-petroleum ether gave 14% yield: mp 95-98°; nmr ($CDCl_3$) showed signals at δ 1.04 (6H, d, isopropyl), 1.45 (3H, d, alanine methyl), 2.08 (3H, s, acetyl), 4.33 (2H, s, $-CH_2Cl$); mass spectrum m/e (relative intensity) 261 (4, P-CH_3), 199 (40, P-$COCH_2Cl$), 114 (25, $CH_3CONHCH(CH_3)$

CO), 86 (100, $CH_3CONHCHCH_3$), 43 (16, CH_3CO).

Anal. Calcd for $C_{12}H_{21}ClN_2O_3$: C, 52.17; H, 7.61; N, 10.14. Found: C, 52.34; H, 7.40; N, 10.01.

N-t-Butyloxycarbonyl-glycylglycyl-L-leucyl Chloromethyl Ketone (Z-Gly-Gly-LeuCH_2Cl). This compound was prepared from Z-Gly-Gly-OH (1.34 g, 5.03 mmoles; prepared from glycylglycine and benzyloxycarbonyl chloride) by a mixed anhydride procedure as previously described. The crude product obtained as an oil was chromatographed on 0.05-0.02 mm Silica Gel G (elution with 4% methanol). Crystallization occurred from ethanol-water; however, when any attempt was made to filter and isolate the crystalline product an oil was obtained. Attempts at crystallization with other solvents indicated that water must be present in order for the crystal structure to form and remain in tact. Therefore, all solvents were removed in vacuo and the product was isolated as an oil (34% yield): nmr ($CDCl_3$) showed signals at δ .94 (6H, d, isopropyl), 3.90 (4H, t, glycine methylenes), 4.26 (2H, s, $-CH_2-Cl$), 5.08 (2H, s, $C_6H_5-CH_2-$), 7.30 (5H, s, C_6H_5-); mass spectrum m/e (relative intensity) 345 (1, P-$2CH_3$-HCl), 271 (82, P-$C_6H_5CH_2-CH_2Cl$), 255 (12, P-$C_6H_5CH_2O-CH_2Cl$), 197 (13, $CONHCH_2CONHCH_2CONHCHCO$), 193 (100), 135 (23, $C_6H_5CH_2OCO$), 107 (6, $C_6H_5CH_2O$), 85 (19, $(CH_3)_2CHCH_2CHNH$), 43 (55, $CH(CH_3)_2$).

Anal. Calcd for $C_{19}H_{26}N_3O_5Cl$: C, 55.40; H, 6.31; N, 10.20. Found: C, 54.14; H, 6.65; N, 10.01.

Acetyl-L-phenylalanyl-glycylglycine Ethyl Ester (Ac-Phe-Gly-Gly-OEt). A mixed carboxylic anhydride was prepared from Ac-Phe-OH (4.14 g, .02 moles; prepared by acylation of L-phenylalanine with acetic anhydride in boiling acetic acid), isobutylchloroformate (2.6 ml, .02 moles), and N-methyl morpholine (2.2 ml, .02 moles) with DMF (25 ml) as the solvent. To this solution was added glycylglycine ethyl ester hydrochloride (3.92 g, .02 moles, prepared by reaction of glycylglycine in absolute ethanol with HCl gas; Greenstein, 1961) and triethylamine (3.0 ml, .02 moles). Following the standard work-up as described above, the product was isolated and recrystallized from hot chloroform yielding 2.4 g (38%): mp 150-151°; nmr (DMSO-d_6) showed signals at δ 1.20 (3H, t, $-CH_2-\underline{CH_3}$), 1.75 (3H, s, acetyl), 2.90 (2H, d, $C_6H_5-\underline{CH_2}$), 4.10 (4H, m, glycine methylene), 7.26 (5H, s, C_6H_5-); mass spectrum m/e (relative intensity) 349 (8, P), 304 (6, P-CH_3CH_2O), 290 (39, $C_6H_5CH{=}CHCONHCH_2CONHCH_2COCH_2CH_3$), 247 (50, $CH_3CONHCH(CH_2C_6H_5)CONHCH_2CO$), 199 (93, $CH_3CONHCH(CH_2)CONHCH_2CONHCH_2$), 131 (58, $C_6H_5CH{=}CHCO$), 120 (100, $C_6H_5CH_2CHNH$), 104 (49, $C_6H_5CH_2CH$), 91 (49, $C_6H_5CH_2$).

Anal. Calcd for $C_{17}H_{23}N_3O_5$: C, 58.45; H, 6.59; N, 12.03. Found: C, 58.27; H, 6.69; N, 11.88.

Acetyl-L-phenylalanyl-glycylglycine (Ac-Phe-Gly-Gly-OH. A suspension of Ac-Phe-Gly-Gly-OEt (1.5 g, 4.3 mmoles) in 20 ml of .25 N NaOH was stirred vigorously for 3 hours at room temperature. The ester gradually dissolved and the

solution was extracted with two 25 ml portions of ethyl acetate. After acidification with HCl to pH 1-2, the water was removed in vacuo to give a white glassy residue. The residue was washed with methanol and filtered. After removal of the methanol, the acid was isolated (not crystallized) and is to be used in the synthesis of Ac-Phe-Gly-Gly-PheCH_2Cl or Ac-Phe-Gly-Gly-LeuCH_2Cl.

Reaction of Subtilisin with Inhibitors

All inhibition experiments were carried out at 30.0°C and at pH 7.0. Stock solutions of the inhibitors were prepared at the appropriate concentrations in dimethoxyethane (DME). The inhibition was started by diluting 0.2 ml of inhibitor solution up to 2.0 ml with 0.1 M phosphate buffer and adding 0.2 ml of subtilisin (contained in pH 7 buffer). The final concentration of enzyme was 5×10^{-6} M and of DME 9 percent. Control experiments showed that this amount of DME did not cause significant change in the specific activity of subtilisin over the time periods of inhibition. Inhibitor concentrations were varied to obtain measurable reaction rates; concentrations were reduced in the case of Z-Gly-Leu-PheCH_2Cl and BOC-Gly-Leu-PheCH_2Cl due to low solubilities of these compounds.

At various time intervals, an aliquot (100 µl) was removed from the inhibition mixture and assayed for residual enzyme activity using N-acetyl-L-tyrosine ethyl ester (Ac-Tyr-OEt). With Ac-Tyr-OEt as substrate, the assay mixtures

contained 0.11 M substrate in 0.1 M KCl containing 8% THF by volume. Titration was performed at pH 7.8 and 30° with .048 N base. The initial velocity was obtained from the slope of the plot of base against time consumption. The pH-stat used was composed of a Radiometer Titrator Type TT 11b, pH Meter Type PHM 26c, Autoburette Type ABU 13, and Recorder Type SBR 2c.

Under the conditions of the assay, the rate of Ac-Tyr-OEt hydrolysis was directly proportional to enzyme concentration. Values of K_{OBS} for each inhibition experiment were calculated from the equation $\ln V = -K_{OBS}t + \text{constant}$ using V and t as input into a Univac 1108 computer operating on a least squares program. The initial rate of substrate hydrolysis at time t is represented by V.

CHAPTER III

RESULTS

Inhibition of subtilisin with a series of peptide chloromethyl ketones was followed as a function of time through at least two half lives. Good pseudo first order kinetics were observed in all cases, and the pseudo first order rate constant (K_{OBS}) was measured from the rate of decrease of residual enzyme activity in the inhibition mixture. A least squares analysis of the data points for each inhibition gave correlation coefficients greater than .95. In Table I the results at 30°, pH 7.0, and 9% 1,2-dimethoxyethane (DME) are shown. Methanol (9%) was used in the case of Ac-Gly-Gly-Ala-PheCH_2Cl due to its low solubility in DME. Comparison of the rate of hydrolysis of Ac-Tyr-OEt in the presence of 9% DME and 9% methanol showed that the activity of subtilisin was essentially the same in either solvent.

Enzyme solutions were prepared fresh each day in 0.1 M phosphate buffer (pH 7.0) to prevent loss of activity from autolysis due to standing in solution overnight. The enzyme concentration varied slightly for individual inhibitions in the range of 5.0 to 5.4 x 10^{-6} M.

Table 1. Data from the Inhibition of Subtilisin with Chloromethyl Ketones. T:30°C; pH 7.0; 9% DME; [E] 5×10^{-6}.

	[I]	K_{OBS}	$K_{OBS}/[I]$	Relative Rates
[b] Ac-Ala-PheCH_2Cl	1.4×10^{-3}	4.92×10^{-4}	3.51×10^{-1}	13.92
	1.4×10^{-3}	5.08×10^{-4}	3.60×10^{-1}	
[c] Ac-Gly-PheCH_2Cl	1.4×10^{-3}	3.70×10^{-5}	2.64×10^{-2}	1.00
	1.4×10^{-3}	3.45×10^{-5}	2.46×10^{-2}	
[d] Ac-Leu-PheCH_2Cl	1.5×10^{-3}	5.30×10^{-5}	3.53×10^{-2}	1.44
	1.5×10^{-3}	5.77×10^{-5}	3.84×10^{-2}	
[b] Ac-Ala-LeuCH_2Cl	1.5×10^{-3}	9.33×10^{-4}	6.22×10^{-1}	25.33
	1.5×10^{-3}	1.00×10^{-3}	6.71×10^{-1}	
[a] Z-Gly-Gly-PheCH_2Cl	2.5×10^{-4}	2.6×10^{-3}	10.60	415.68
[a] Z-Gly-Gly-LeuCH_2Cl	6.9×10^{-5}	1.09×10^{-3}	15.84	651.37
	6.9×10^{-5}	1.20×10^{-3}	17.39	
[b] BOC-Gly-Leu-PheCH_2Cl	2.5×10^{-4}	5.86×10^{-4}	2.34	85.09
	2.5×10^{-4}	5.00×10^{-4}	2.01	

[a]	BOC-Ala-Gly-PheCH$_2$Cl	2.49×10^{-4} 2.49×10^{-4}	2.55×10^{-3} 2.97×10^{-3}	10.25 11.94	434.90
[a]	Z-Gly-Leu-PheCH$_2$Cl	1.2×10^{-4} 1.2×10^{-4}	4.22×10^{-4} 4.37×10^{-4}	3.35 3.47	133.72
[a]	Ac-Ala-Gly-PheCH$_2$Cl	2.49×10^{-4}	1.07×10^{-3}	4.30	168.62
[d]	Ac-Gly-Gly-Ala-PheCH$_2$Cl	2.50×10^{-4}	2.10×10^{-3}	8.78	344.31
[a]	Ac-Ala-Ala-Pro-AlaCH$_2$Cl	2.70×10^{-4}	1.63×10^{-3}	6.05	237.25

[a] 5×10^{-6} [E]

[b] 5.4×10^{-6} [E]

[c] 5.6×10^{-6} [E]

[d] 5.8×10^{-6} [E]

Kinetics of Inactivation

A simple bimolecular reaction between an enzyme and inhibitor may be represented by the chemical equation

$$E + I \xrightarrow{K_2} E - I$$

and the second order rate equation for this reaction is given by

$$v = \frac{d[E-I]}{dt} = - \frac{d[E]}{dt} = K_2[E]\ [I] \qquad (1)$$

However, under the special conditions when the inhibitor concentration is much greater than the enzyme concentration, [I] is essentially unchanging and equation (1) reduces to equation (2).

$$v = - \frac{d[E]}{dt} = K_{OBS}\ [E] \qquad (2)$$

or in the integrated form

$$\ln\ [E]/[E_o] = -K_{OBS}\ t \qquad (3)$$

where [E] is the enzyme concentration at time t, $[E_o]$ is the initial enzyme concentration at time O, and K_{OBS} is the experimentally determined rate constant obtained from the Ac-Tyr-OEt assay of residual enzyme activity in the inhibition mixture at specified time intervals. Thus, when [I] >> $[E_o]$, the reaction is bimolecular, but exhibits pseudo first order kinetics with the observed first order rate constant (K_{OBS}, in dimensions of reciprocal time) equal to

the true second order rate constant multiplied by the inhibitor concentration (equation 4).

$$K_{OBS} = K_2[I] \tag{4}$$

If both sides of equation (4) are divided by [I], $K_{OBS}/[I]$ has the dimensions of concentration $^{-1}$ time $^{-1}$ and is equal to the second order rate constant (equation 5).

$$\frac{K_{OBS}}{[I]} = K_2 \tag{5}$$

The kinetic equations described above are generally applicable to enzyme catalyzed reactions. However, the kinetics of inhibition of certain enzymes by irreversible inhibitors reveals the presence of a reversible enzyme-inhibitor complex preceeding covalent bond formation. Evidence for this behavior has been demonstrated in the inhibition of cholinesterase by certain methanesulfonic acid esters (Kitz and Wilson, 1962), trypsin by Tos-LysCH_2Cl (Shaw and Glover, 1970), carboxypeptidase A by N-ethyl-5-phenylisooxazolium-3'-sulfonate (Petra, 1971), and chymotrypsin by $MeSO_2$-DL-PheCH_2Cl (Kumar and Hein, 1970). Conclusive evidence has not been obtained to show that inhibition of subtilisin by peptide chloromethyl ketones occurs via an intermediate complex. However, the finding that subtilisin will not react with Tos-PheCH_2Cl (Smith et al., 1966) but will react with Z-PheCH_2Cl (Morihara and Oka, 1970)

and Z-PheCH_2Br (Shaw and Ruscica, 1968) seems to indicate that prior to covalent bond formation a prerequisite enzyme-inhibitor complex is formed and the inhibitor must be structured so that binding to the enzyme is favorable. Consequently, experimentally determined K_{OBS} values must be related to the kinetic parameters associated with an intermediate complex.

The irreversible reaction of an active site directed inhibitor involving an intermediate complex may be represented by the overall reaction sequence

$$E + I \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E \cdot I \xrightarrow{k_3} E - I$$

where E·I is the reversible noncovalently bound complex of the enzyme with the inhibitor and E-I represents the inactive enzyme with inhibitor irreversibly bound via a covalent linkage. The following equations apply:

$$[E^\circ] = [E] + [E\text{-}I] \qquad (6)$$

where $[E^\circ]$ is the initial concentration of enzyme and at any time is equal to all species involving enzyme in this inhibitor solution; and

$$[E] = [E] + [E \cdot I] \qquad (7)$$

where $[E]$ is the concentration of free and noncovalently bound enzyme. Since k_3 is the limiting rate of inactivation,

the dissociation constant of the E·I complex may be defined as

$$K_I = k_2/k_1 = [E]\ [I]/[E\cdot I] \tag{8}$$

Substituting [E] = [E] - [E·I] from equation (7) into equation (8) and solving for [E·I] gives equation (9).

$$[E\cdot I] = \frac{[E]\ [I]}{K_I + [I]} \tag{9}$$

If the inhibitor concentration is sufficiently greater than the total enzyme concentration, the decrease in [E] in the inhibition mixture is dependent on the concentration of the enzyme inhibitor complex.

$$-\frac{d[E]}{dt} = k_3[E\cdot I] \tag{10}$$

The concentration of both free and noncovalently bound enzyme represented by [E] is measured by the assay due to the dilution of the reversible enzyme inhibitor complex in the assay mixture.

Substitution of equation (9) into equation (10) gives

$$-\frac{d[E]}{dt} = k_3[E]\,[I]/K_I+[I] \tag{11}$$

and rearranging (Kitz and Wilson, 1962).

$$-\frac{d[E]}{dt} = k_3\,\frac{1}{K_I/[I]+1} \tag{12}$$

Since [I] is essentially constant, equation (12) may be integrated to give equation (13).

$$\ln [E]/[E^\circ] = -k_3 t \frac{1}{K_I/[I]+1} \quad (13)$$

Equation (13) is related to the pseudo first order rate constant (K_{OBS}) in equation (3). Thus,

$$-K_{OBS}\ t = -k_3 t \frac{1}{K_I/[I]+1} \quad (14)$$

which reduces to

$$K_{OBS} = \frac{k_3}{K_I/[I]+1} \quad (15)$$

This equation may be expressed in its reciprocal form (equation 16).

$$\frac{1}{K_{OBS}} = \frac{K_I}{k_3[I]} + \frac{1}{k_3} \quad (16)$$

This derivation shows that the decrease in free enzyme concentration [*E*] in the inhibition mixture follows pseudo first order kinetics at any fixed value of [I] if [I] [E] and the observed first order rate constant is represented by equation (15) or its reciprocal form, equation (16). Therefore, in general K_{OBS} is not constant with changing values of [I], and the exact meaning of K_{OBS} depends upon the relative magnitude of the inhibition concentration and the binding constant (K_I) of the enzyme-inhibitor complex. If $[I] \cong K_I$, equation (15) defines K_{OBS}, and K_I and k_3 may be evaluated by use of a double reciprocal plot and equation (16). If $[I] >> K_I$, equation (15) reduces to

$$K_{OBS} = k_3 \tag{17}$$

and the K_{OBS} is a measure of the rate of covalent bond formation in the enzyme-inhibitor complex. If $[I] << K_I$, equation (15) reduces to

$$\frac{K_{OBS}}{[I]} = \frac{k_3}{K_I} \tag{18}$$

and the second order rate constant $K_{OBS}/[I]$ reflects both the influence of K_I and k_3.

Since inhibitor binding constants cannot be obtained from the data presented in Table I, the specific equation defining K_{OBS} cannot be conclusively determined. However, Glazer (1967) measured K_I's for a series of competitive inhibitors of the hydrolysis of Ac-Tyr-OEt by subtilisin and found a range of values from 30 to 340 mM. Also, Morihara et al. (1970) measured K_m's of a series of synthetic di-, tri-, and tetrapeptide amides and found a range of 5.7 to 80 mM (K_m, in certain instances, is a binding constant for substrates, analogous to the binding constant K_I for inhibitors). Furthermore, Kurachi, Powers and Wilcox (1971) placed a lower limit of about 8 mM for the K_I's of inhibition of Chymotrypsin A by peptide chloromethyl ketones. Since chymotrypsin and subtilisin are homologous proteases and are similar in their mode of reaction with peptide chloromethyl ketones, binding constants are likely to be in the same general range (Robertus et al., 1972). The range of

inhibitor concentrations from Table I is .069 to 1.5 mM. Since the minimum previously determined K_I value was 5.7 mM, it is likely that K_I for these peptide chloromethyl ketones is greater than the inhibitor concentrations used and

$$\frac{K_{OBS}}{[I]} = \frac{k_3}{K_I} \qquad (18)$$

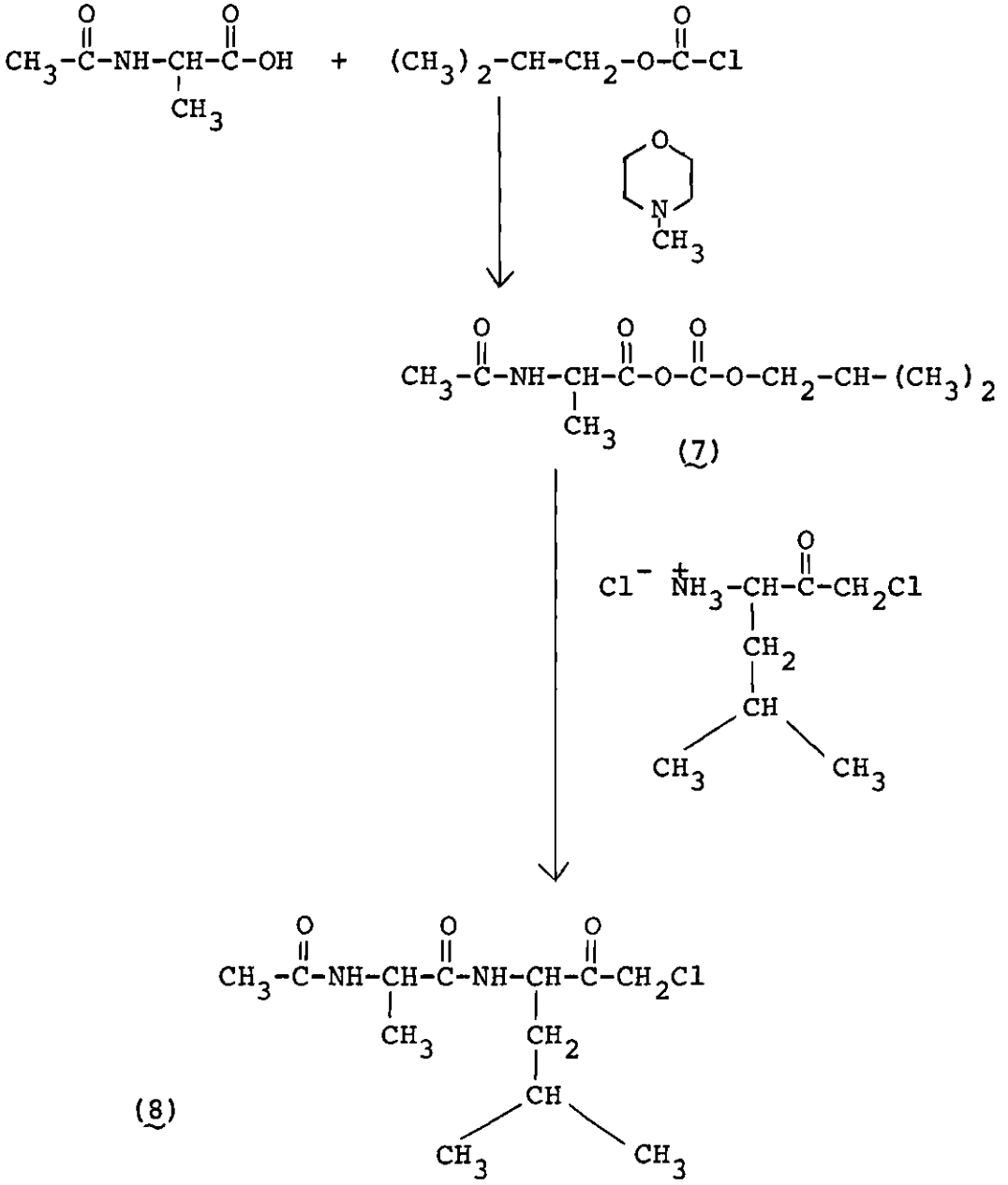
Thus is seems reasonable to assume that the second order rate constant ($K_{OBS}/[I]$) reflects both the influence of K_I and k_3 according to equation (18).

CHAPTER IV

Discussion

Synthesis of Inhibitors

The majority of the amino acid couplings were done via the mixed carboxylic anhydride procedure (Anderson et al., 1952). Using this procedure, the carboxyl group of an amino acid residue is activated under anhydrous conditions by reaction with isobutylchloroformate (or isovaleryl chloride) to give the mixed anhydride intermediate (7). The anhydride is then reacted with the free amine of another residue giving the dipeptide (8). An example of this method is shown on the following page in the preparation of the leucine chloromethyl ketone, Ac-Ala-LeuCH_2Cl.



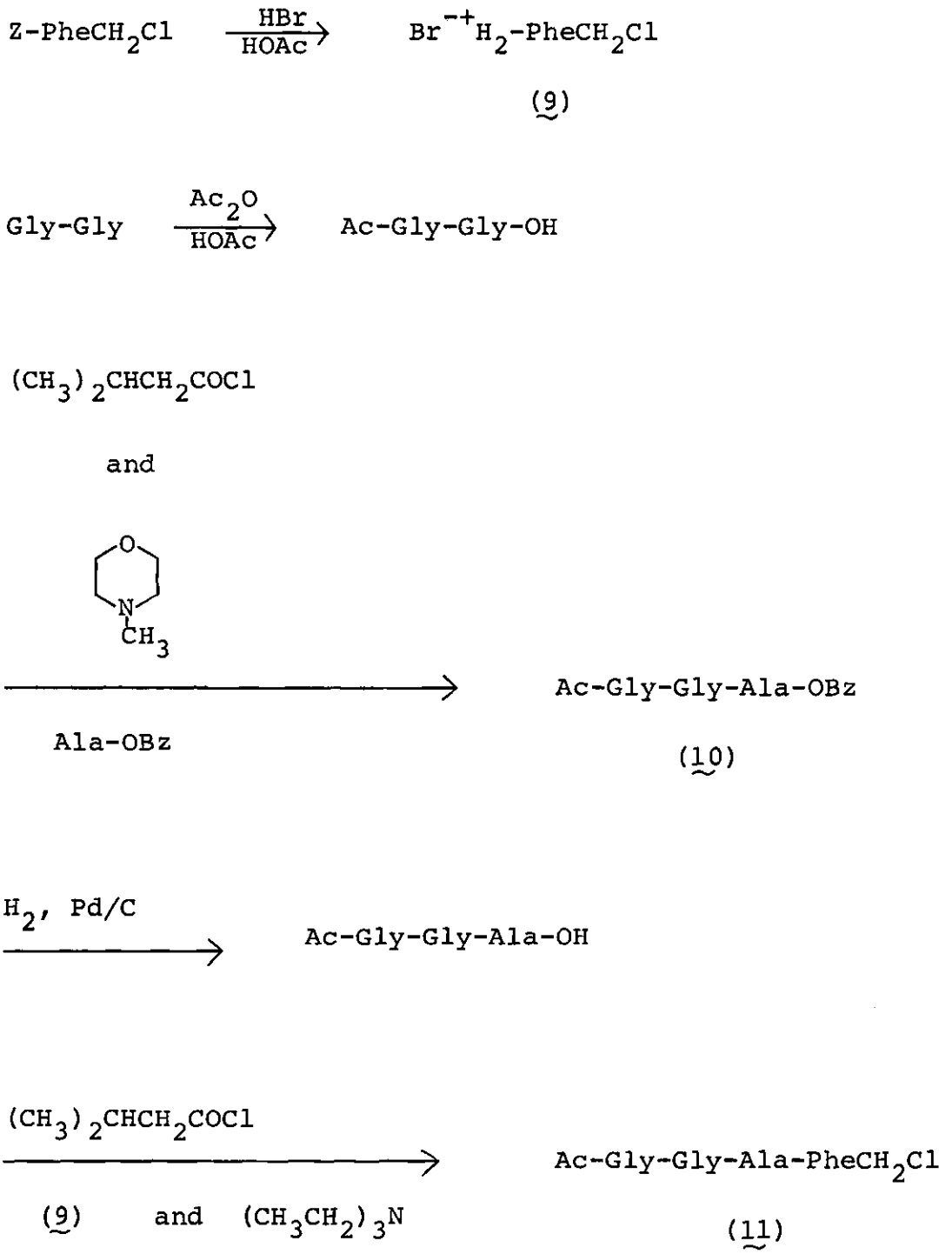
(7)

(8)

Scheme 1

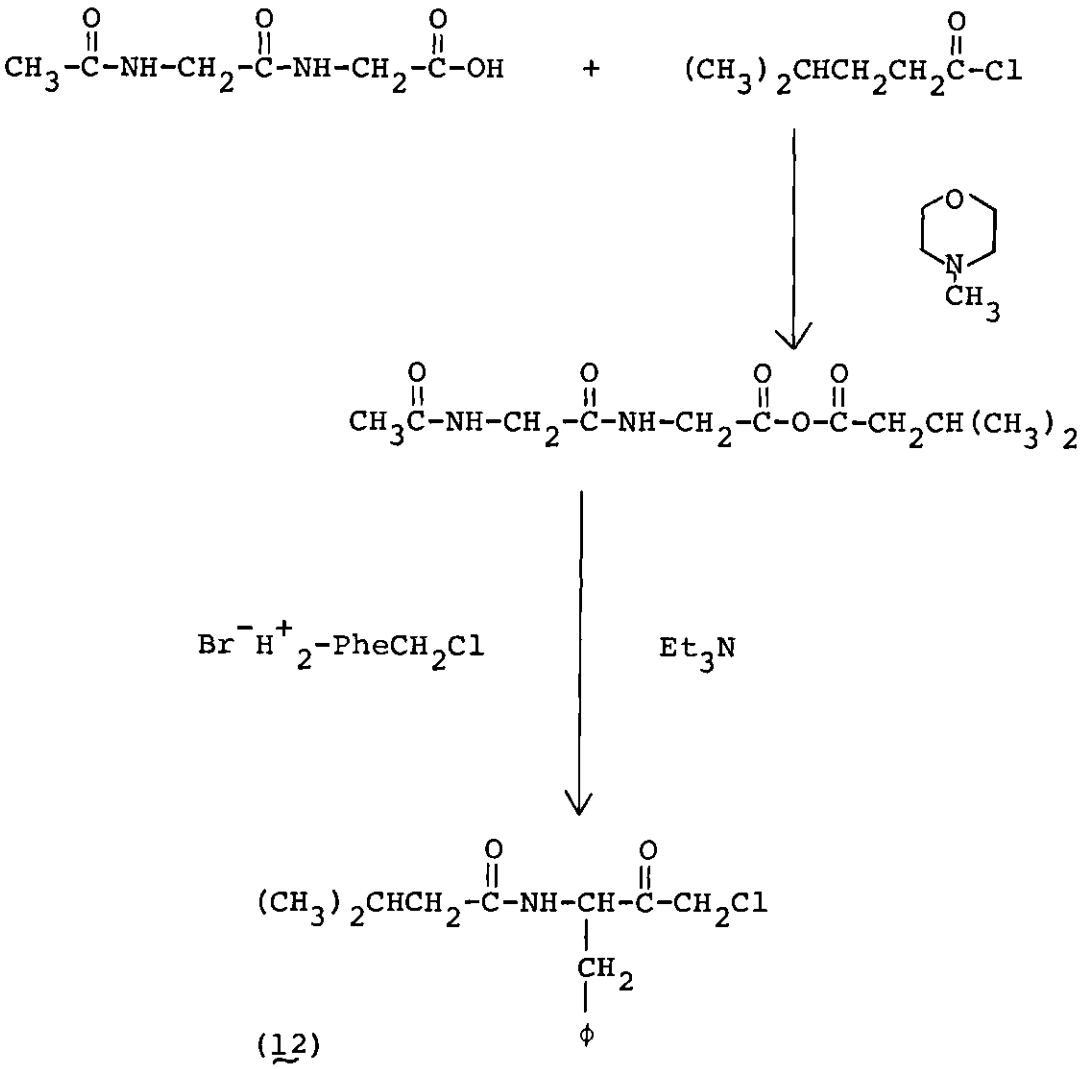
The carbonyl carbon atom of the mixed anhydride most susceptible to attack by a nucleophile should be the one with the lowest electron density and least steric hindrance. Consequently, in the example shown previously, attack occurred at the carbonyl of alanine giving the dipeptide as the product.

The first objective in the synthetic work was the preparation of a tetrapeptide L-phenylalanine chloromethyl ketone. The scheme shown on the next page shows the synthetic procedure used to obtain this compound (Ac-Gly-Gly-Ala-PheCH_2Cl).


Z-PheCH2Cl
HBr
HOAc
Br-+H2-PheCH2Cl
(9)
Gly-Gly
Ac2O
HOAc
Ac-Gly-Gly-OH
(CH3)2CHCH2COCl
and
O
N
CH3
Ala-OBz
Ac-Gly-Gly-Ala-OBz
(10)
H2, Pd/C
Ac-Gly-Gly-Ala-OH
(CH3)2CHCH2COCl
(9) and (CH3CH2)3N
Ac-Gly-Gly-Ala-PheCH2Cl
(11)

Scheme 2

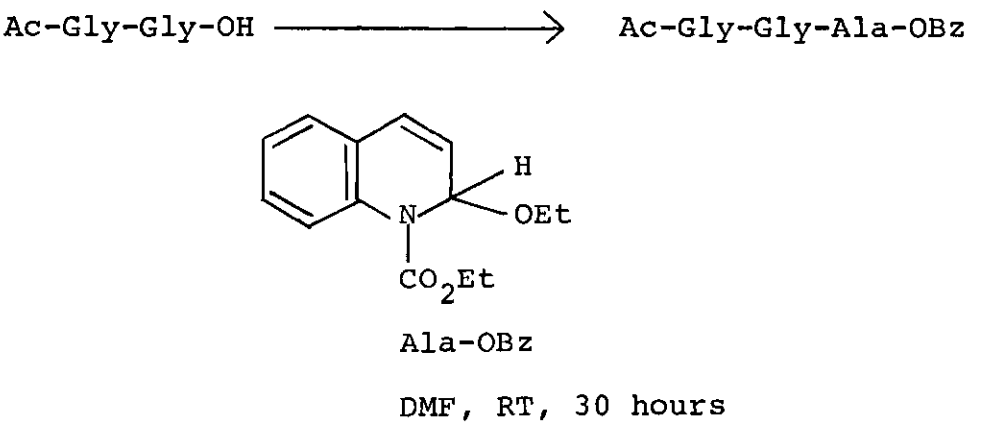
The tripeptide ester (Ac-Gly-Gly-Ala-OBz, 10) was obtained in 60% yield via the mixed anhydride procedure using the iso-valeryl chloride. No attempt was made to prepare this compound using iso-butylchloroformate. On the other hand, the tetrapeptide chloroketone (Ac-Gly-Gly-Ala-PheCH_2Cl, 11) was prepared using both of these reagents. The yield using iso-butylchloroformate was 44% and no byproduct was isolated. However, using iso-valeryl chloride the yield of compound (11) was 30% and a significant amount (yield not determined) of iso-valeryl phenylalanine chloromethyl ketone (12) was obtained resulting from attack by the amine moiety of phenylalanine chloromethyl ketone at the wrong carbonyl of the mixed anhydride as shown on the next page.



(12)

Scheme 3

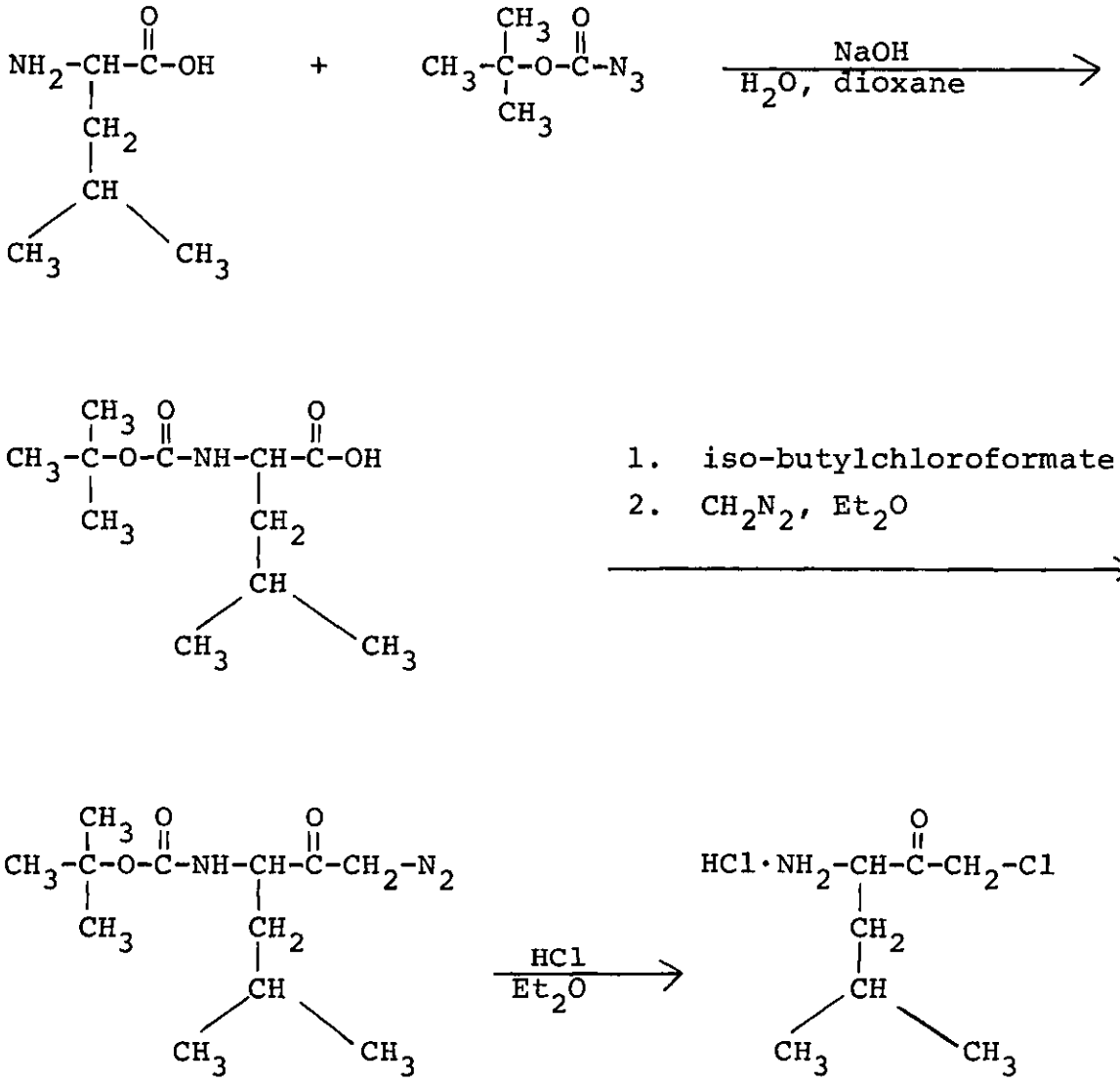
Compound (10) was also obtained in 44% yield using EEDQ. This procedure is shown schematically on the next page:

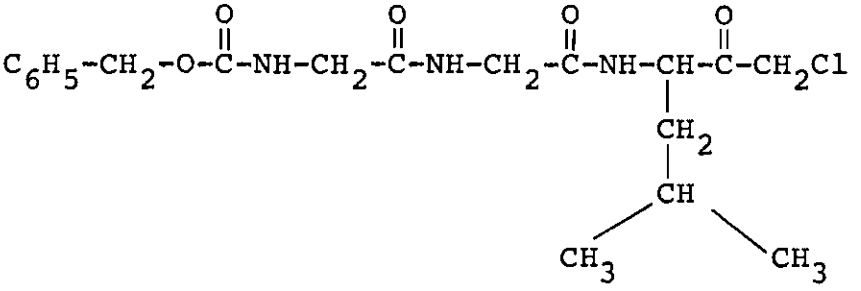


Scheme 4

The procedure and work-up were very simple but the yield was relatively low. No attempt was made to vary conditions in order to improve the yield. The preparation of compound (11) was also attempted via the EEDQ procedure, but no product was isolated.

The key intermediate in the synthesis of the leucine chloromethyl ketones was L-leucine chloromethyl ketone hydrochloride obtained from the reaction of gaseous hydrochloric acid and t-butyloxycarbonyl-L-leucine diazoketone. The scheme for the preparation of Z-Gly-Gly-LeuCH_2Cl is shown on the next page.


NaOH
H2O, dioxane
1. iso-butylchloroformate
2. CH2N2, Et2O
HCl
Et2O
Z-Gly-Gly-OH,
mixed anhydride
(iso-butylchloroformate)



Scheme 5

An initial attempt was made to prepare the hydrobromide of leucine chloromethyl ketone by deblocking Z-LeuCH_2Cl with HBr in acetic acid. All attempts failed and the hydrobromide was never isolated.

The compound Ac-Phe-Gly-Gly-OEt was obtained by the mixed anhydride procedure analogous to those described above. The ester was hydrolyzed and a crude oil was isolated. Due to lack of time, no spectra were obtained on the oil, but the compound was believed to be the acid due to the release of CO_2 when the oil was dissolved in a sodium bicarbonate solution. However, attempts to couple the acid with phenylalanine chloromethyl ketone were unsuccessful and the synthesis was not completed.

Structure and Reactivity of Inhibitors

An x-ray crystallographic study of subtilisin inhibited by four peptide chloromethyl ketones (Z-Ala-PheCH_2Cl, Z-Ala-Gly-PheCH_2Cl, Z-Gly-Gly-PheCH_2Cl, and Ac-Ala-Gly-

PheCH$_2$Cl) has shown that these inhibitors are bound to the enzyme via a covalent linkage between the imidiazole ring of His-64 and the methylene group of the chloromethyl ketone moiety (Robertus et al., 1972). The benzyl group of the phenylalanine residue lies in a hydrophobic crevice and the longest inhibitor formed three hydrogen bonds with the enzyme, producing a β-sheet structure with an extended segment of backbone chain in the enzyme consisting of residues Ser-125, Leu-126, and Gly-127. The study presented in this thesis was undertaken to determine whether the solution reactivity of a series of peptide chloromethyl ketones could be correlated with the crystal structures of these inhibited subtilisin derivatives and with kinetic data from recent experiments with polypeptide substrates (Morihara et al. 1969, 1970).

The rate of irreversible reaction of a site specific inhibitor with an enzyme should depend upon the amount of enzyme-inhibitor complex present at equilibrium and the rate limiting reaction of the bound inhibitor with the enzyme to form the inactivated enzyme. The steric arrangement of the inhibitor and enzyme in the complex and electronic effects would certainly influence covalent bond formation. On the other hand, the dissociation constant K_I of the E·I complex is determined primarily by an entropy effect. An increase in entropy occurs upon binding the inhibitor to the enzyme due to breaking up of some water structure around

the water separated inhibitor and enzyme relative to the E·I complex. A structural change in the inhibitor could effect the extent of binding to the enzyme (measured by K_I), the stereoelectronic relationship between inhibitor and enzyme in the E·I complex (measured by k_3) or both (Kurachi, Powers and Wilcox, 1972). Shaw and Glover (1970) have shown that for inhibition of trypsin with certain chloromethyl ketones, the affinity of the enzyme for the inhibitor is not necessarily related to the rapidity of inactivation. In this thesis, these two effects were not determined and due to the low inhibitor concentrations used, the second order rate constants ($K_{OBS}/[I]$) used in the discussion reflect the influence of both K_I and k_3.

As proposed from the x-ray results, a binding scheme for one of the peptide chloromethyl ketones used in this study is shown in Figure 4. This illustration is based on the crystallographically determined model of subtilisin inhibited with Z-Ala-Gly-PheCH_2Cl. In the following discussion of the reactivity of various inhibitors in terms of this model, the notations previously described (Chapter I, Figure 2) for peptide binding subsites for proteolytic enzymes will be used.

Examination of the data in Table I (Chapter III) shows that there is a spread of a factor of 651 between the least reactive inhibitor (Ac-Gly-PheCH_2Cl) and the most reactive (Z-Gly-Gly-LeuCH_2Cl). In general, the larger the number

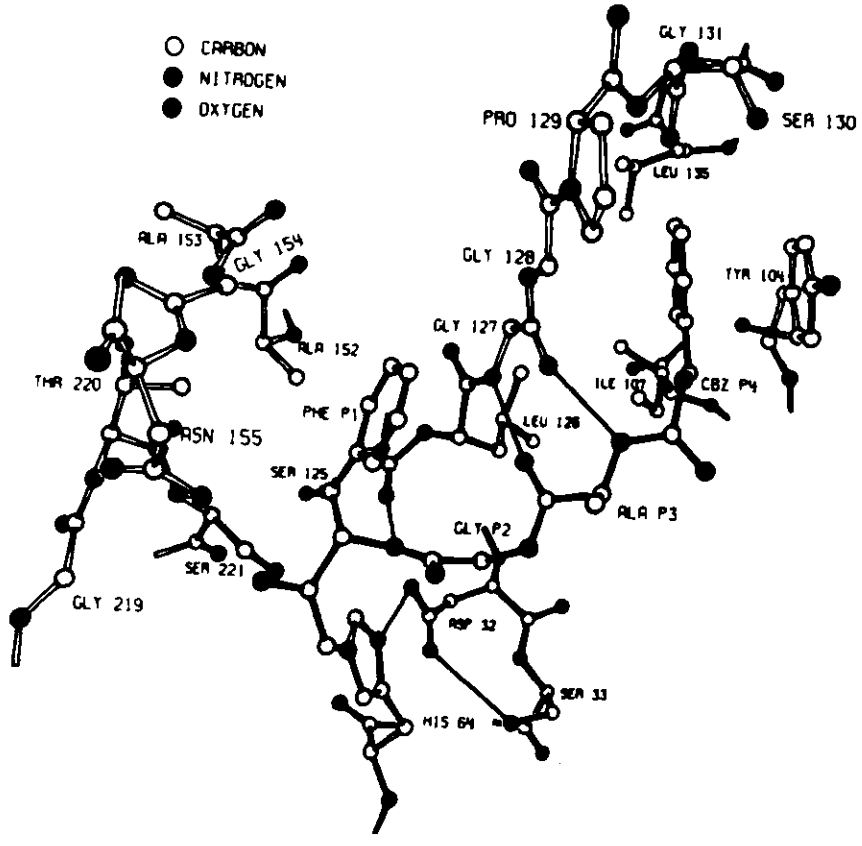


Figure 4. Crystallographic Model of Subtilisin Inhibited with Z-Ala-Gly-PheCH_2Cl.

of positive interactions between a given inhibitor and the enzyme, the larger the second order rate constant. This phenomena has previously been observed with peptide amide substrates (Morihara et al. 1970) as shown by the relative rate constants below in Table 2.

Table 2. Comparison of Relative Rates of Di- and Tetra-peptides Inhibitors and Substrates.

	K_{cat}
Z-Gly-Leu-NH_2	1.00
Z-Ala-Gly-Gly-Leu-NH_2	86.78
	$K_{OBS}/[I]$
Ac-Ala-PheCH_2Cl	1.00
Ac-Gly-Gly-Ala-PheCH_2Cl	24.73

It has been known for some time that the primary specificity of subtilisin is determined by the residue at the P_1 subsite in an inhibitor or substrate, and subtilisin shows a preference for amino acid residues in this position with aromatic or apolar side chains of the L-configuration.

This observation is readily explained by the presence of a hydrophobic binding cleft at site S_1. Although the proposed substrate binding model from the x-ray study was based exclusively upon experiments with a phenylalanine residue at P_1, the authors postulated that an identical crystallographic model could be constructed with a leucine residue at P_1. The following relative rates in Table 3 indicate that subtilisin shows a preference for leucine at P_1.

Table 3. Relative Rates of Inhibitors Indicating Preference for Leucine Residue at P_1.

Ac-Ala-LeuCH_2Cl	(25)
Ac-Ala-PheCH_2Cl	(14)
Z-Gly-Gly-LeuCH_2Cl	(651)
Z-Gly-Gly-PheCH_2Cl	(416)

Also, Morihara and Tsuzuki (1969) found that subtilisin showed a preference for leucine over phenylalanine at P_1, as shown in Table 4. Future x-ray work on subtilisin inhibited with Z-Gly-Gly-LeuCH_2Cl is planned and will hopefully explain subtilisin's preference for leucine at P_1.

Table 4. Relative Rates of Hydrolysis of Ester Substrates Indicating a Preference for Leucine Residue at P_1.

	K_{cat} (sec^{-1})
Ac-Leu-OMe	57.5
Ac-Phe-OMe	30.6

A somewhat surprising result was that alanine at P_1 showed a significant relative reaction rate. The relative rate of Ac-Ala-Ala-Pro-AlaCH_2Cl was 237.

This apparent broad specificity at S_1 for these tri- and tetra- peptide inhibitors is indicative of the importance of the other subsites (S_2-S_4) in binding. It would be of interest to determine the reactivity of a dipeptide inhibitor with alanine at P_1 in comparison with the other dipeptides listed in Table 1.

Strong evidence supporting the postulation that the solution reactivity of peptide chloromethyl ketones and of substrates can be correlated with the crystallographically based model is the effect on reactivity that an alanine residue has at P_2 of a substrate or inhibitor. The $K_{OBS}/[I]$ value of Ac-Ala-PheCH_2Cl is greater than Ac-Gly-PheCH_2Cl and Ac-Leu-PheCH_2Cl by a factor of about 13, although they have the same structure except for the P_2 residue. These

results agree beautifully with kinetic data of Morihara et al. (1970) on the hydrolysis of peptide substrates by subtilisin. A comparison of the relative rates of the inhibitor and amide substrates is shown in Table 5.

Table 5. Comparison of Relative Rates of Inhibitors and Substrates with Alanine Residue at P_2.

	$K_{OBS}/[I]$		K_{cat}
Ac-Ala-PheCH_2Cl	(13.92)	Z-Ala-LeuNH_2	(14.00)
Ac-Leu-PheCH_2Cl	(1.44)	Z-Leu-LeuNH_2	(2.40)
Ac-Gly-PheCH_2Cl	(1.00)	Z-Gly-LeuNH_2	(1.00)

These observations are consistent with the crystallographic results which show that the methyl side chain of alanine at P_2 makes van der Waals contact with Leu-96 and the side chain of His-64. It was presumed that these contacts allow optimum orientation of the substrate in the active site region and a larger side chain at P_2 would disrupt this alignment, while glycine at this position would allow greater freedom of motion and hence retard proper alignment of the inhibitor.

The reason that leucine is slightly faster than glycine at P_2 in the above series is not immediately obvious, and these results are apparently anomolous with the finding that Z-Gly-Gly-PheCH_2Cl reacts almost 200 times faster than Z-Gly-Leu-PheCH_2Cl. It is possible that retardation of alignment resulting from freedom of motion with glycine at P_2 in the dipeptide series, is cancelled due to extension of the peptide chain to S_3 and subsequent stabilization by the addition of two hydrogen bonds. On the other hand, the unfavorable steric interactions resulting from leucine at P_2 are neutralized to a much lesser degree by extension of the peptide chain.

The P_3-S_3 interaction involves two additional hydrogen bonds and as expected, the kinetic data in Table 1 indicates that tripeptides are better inhibitors than dipeptides. For example, extension of the dipeptide inhibitor Ac-Gly-PheCH_2Cl to Z-Gly-Gly-PheCH_2Cl or BOC-Ala-Gly-PheCH_2Cl increases the second order rate constant by a factor of over 400. Also, extension of Ac-Leu-PheCH_2Cl to Z-Gly-Leu-PheCH_2Cl increases $K_{OBS}/[I]$ by a factor of about 150.

It is clear from the kinetic data of Morihara et al. (1970) that subsite S_4 exerts an influence on substrate-enzyme reaction. These authors found that the tetrapeptide substrate Z-Ala-Gly-Gly-LeuNH_2 was hydrolyzed over 2.5 times faster than the tripeptide substrate Z-Gly-Gly-LeuNH_2. Very little can be said about S_4 from the x-ray results since a

tetrapeptide was not studied. A benzyloxycarbonyl or acetyl blocking group was located at P_4 and determination of the exact location of these groups was hindered for reasons which will not be discussed. However, the electron density maps from the enzyme inhibited with Z-Ala-PheCH_2Cl showed that a large depression in the sirface of the enzyme existed in the vicinity of S_4 which could possibly represent a second hydrophobic crevice for binding P_4 residues. This suggests that S_4 may be designed with structural specifications similar to S_1. The data in Table 1 shows that the fastest inhibitors are Z-Gly-Gly-LeuCH_2Cl, Z-Gly-Gly-PheCH_2Cl, and BOC-Ala-Gly-PheCH_2Cl. All have aromatic or apolar systems at P_4. Although these are not amino acid residues, it is possible that good binding occurs due to the P_4-S_4 interactions described above. These three inhibitors are faster than the tetrapeptide Ac-Gly-Gly-Ala-PheCH_2Cl. Replacement of the BOC groupd in BOC-Ala-Gly-PheCH_2Cl with an acetyl group decreases the second order rate constant approximately 38 percent.

Crystallographic work is planned on subtilisin inhibited with Ac-Gly-Gly-Ala-PheCH_2Cl which should give indication of the interaction at S_4 with an amino acid residue at P_4. Also, synthesis of a tetrapeptide with a phenylalanine residue at P_4 has been started to determine the effect of an aromatic side chain in the postulated hydrophobic cavity at S_4.

In conclusion, the kinetic results presented in this thesis are consistent with the model illustrated in Figure 4 and show that the magnitude of the second order rate constant $K_{OBS}/[I]$ is dependent on the number of positive interactions between the enzyme and inhibitor. These rates are also correlated with substrate specificity of subtilisin and support the hypothesis that the productive binding mode of this enzyme with chloromethyl ketone inhibitors which leads to alkylation of His-64 is similar to the productive binding mode of substrates leading to acylation of Ser-221.

CHAPTER V

RECOMMENDATIONS

The structure analysis by x-ray of Ac-Gly-Gly-Ala-PheCH_2Cl would provide exact information regarding interactions at the fourth subsite; this work is planned in the near future. Also, an x-ray study on Z-Gly-Gly-LeuCH_2Cl is planned to observe interactions in the S_1 subsite with a leucine residue at P_1.

The synthesis of Ac-Phe-Gly-Gly-PheCH_2Cl or Ac-Phe-Gly-Gly-LeuCH_2Cl has been started in order to observe via crystallography possible hydrophobic interactions at S_4 with the phenylalanine side chain at P_4.

A kinetic study to determine the variation in the ratio $K_{OBS}/[I]$ with inhibitor concentration would provide information for determining the exact equation defining the pseudo first order rate constants, K_{OBS}, obtained in this thesis.

References

Alden, R., Wright, C., and Kraut, J. (1970) Phil. Trans. Roy. Soc. (London) B257, 119.

Anderson, G.W., Blodinger, J., Welcher, A.D. (1952) J. Amer. Chem. Soc. 74, 5309.

Birktoft, J., Blow, D.M., Henderson, R. and Steitz, T.A. (1970) Phil. Trans. Roy. Soc. (London) B257, 67.

Blow, D.M., Birktoft, J. and Hartley, B.S. (1969) Nature 221, 337.

Hartley, B.S. (1960) Annu. Rev. Biochem. 29, 45.

Kitz, R. and Wilson,I. (1962) J. Biol. Chem. 257, 3245.

Kraut, J. (1971) "The Enzymes," Vol. III, 3rd Edition, Chapter 15.

Kumar, S. and Hein, G. (1970) Biochem. Biophys. Acta 206, 404.

Kurachi, K., Powers, J., and Wilcox, P. Biochemistry (in press).

Lehninger, A. (1970) "Biochemistry," Worth Publishers, New York, p. 169.

Linderstrom-Lang, K. and Ottensen, M. (1947) Nature 159, 807.

Markland, F.S., Shaw, E., and Smith, E.L. (1968) Proc. Natl. Acad. Sci. U.S. 61, 1440.

Matsubara, H., Hagihara, B., Nakai, M., Komaki, T., Yonetani, T., and Okanuki, K. (1958) J. Biochem. (Tokyo) 45, 251.

Moore, J.A. and Reed, D.E. (1961) Org. Syn. 41, 16.

Morihara, K. and Oka, T. (1970) Arch. Biochem. Biophys. 138, 526-531.

Morihara, K. and Oka, T. and Tsuzuki, H. (1970) Arch. Biochem. Biophys. 138, 515-525.

Morihara, K. and Tsuzuki, H. (1969) Arch. Biochem. Biophys. 129, 620.

Noller, H.F. and Bernhard, S.A. (1965) Biochemistry 4, 1118.

Okanuki, K., Matsubara, H., Nishimura, S. and Hagihara, B. (1956) J. Biochem. (Tokyo) 43, 857.

Oosterbahn, R.A. and Cohen, J.A. (1964) "Structure and Activity of Enzymes," (Goodwin, T.W., Harris, J.I. and Hartley, B.S., Editors) New York p. 87.

Petra, P.H. (1971) Biochemistry 10, 3163.

Polgar, L. and Bender, M.L. (1966) J. Amer. Chem. Soc. 88, 3153.

Polgar, L. and Bender, M.L. (1967) Biochemistry 6, 610.

Powers, J.C. and Wilcox, P.E. (1970) J. Amer. Chem. Soc. 92, 1782.

Roberts, J.D. (1961) Org. Syn. 41, 16.

Robertus, J., Alden, R., Birktoft, J., Kraut, J., Powers, J. and Wilcox, P. (1972) Biochemistry 11, 2439.

Schecter, I. and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157.

Schnabel, E. (1967) Justus Liebigs Ann. Chem. 702, 188.

Shaw, E. and Glover, G. (1970) Arch. Biochem. Biophys. 139, 298.

Shaw, E. and Ruscica, J. (1968) J. Biol. Chem. 243, 6312.

Smith, E.L., Markland, F.S., Kasper, C.B., Delange, R.J., Landon, M. and Evans, W.H. (1966) J. Biol. Chem. 241, 5974.